Year: 2017

Characterizing change points and continuous transitions in movement behaviours using wavelet decomposition

Soleymani, Ali ; Pennekamp, Frank ; Dodge, Somayeh ; Weibel, Robert

DOI: https://doi.org/10.1111/2041-210X.12755

Originally published at:
Soleymani, Ali; Pennekamp, Frank; Dodge, Somayeh; Weibel, Robert (2017). Characterizing change points and continuous transitions in movement behaviours using wavelet decomposition. Methods in Ecology and Evolution, 8(9):1113-1123.
DOI: https://doi.org/10.1111/2041-210X.12755

Appendix S1

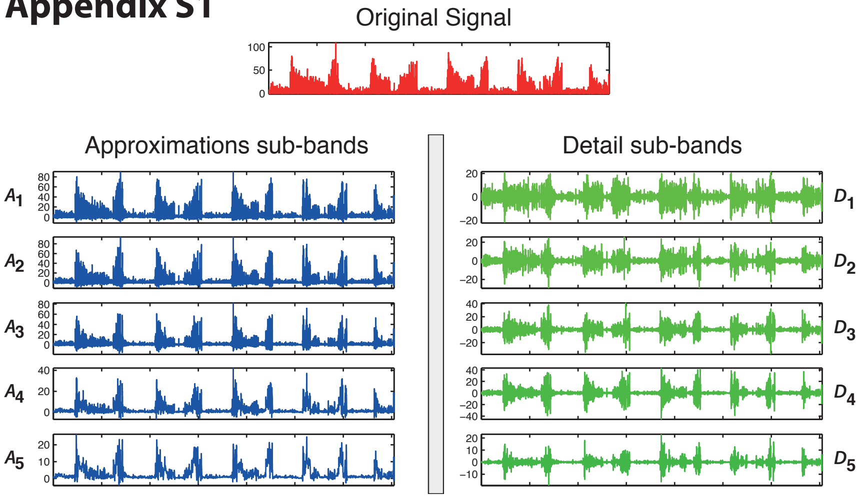

Figure S1. Example of decomposition of a signal using DWT. The signal is passed through low-pass and high-pass filters to generate approximation and detail sub-bands. *Aj* represents the approximation sub-band at the *j*th level of decomposition using a low-pass filter and *Dj* represents the corresponding detail sub-band at the *j*th level obtained through the high-pass filter.

Appendix S2

Detailed implementation of DWT-based segmentation method

1. Background

In the following, a step-by-step guideline for the application of the DWT-based method will be provided. An overview of these steps is given in Figure S2a.

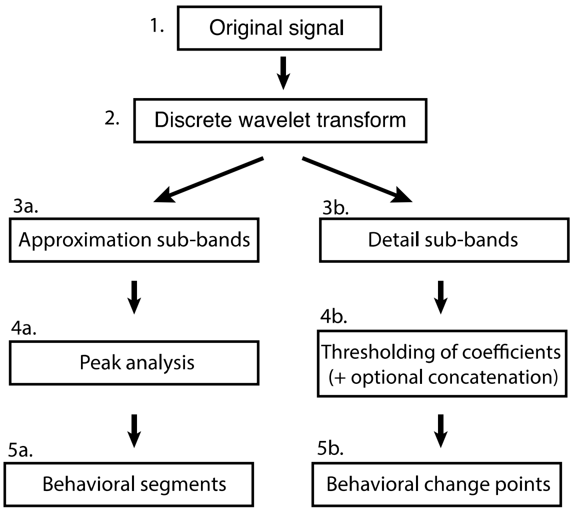

Figure S2a: Flowchart of the DWT segmentation method

In Step 1, the movement parameter profile (i.e. time series of the chosen movement parameter; speed in our example) is calculated from the raw movement data, acting as the input for the DWT transform. In Step 2, the movement signal is decomposed by the DWT. This transformation yields two outputs: low frequency approximation sub-bands (3a) and high frequency detail sub-bands (3b). These two resulting sets of sub-bands will be used for different purposes in the next stage and therefore the procedure bifurcates into two different branches, along which different analyses will be performed. Peak analysis will be performed on approximation sub-bands (Step 4a) in order to obtain behavioural segments (Step 5a), while thresholding of detail sub-bands (Step 4b) will be used to detect behavioural change points (Step 5b).

2. Worked example for the implementation of the DWT segmentation

We will use the movement trajectory of turkey vulture individual *Leo* to show how the method is applied in practice and how the parameters are set at each stage. The movement trajectory of Leo is shown in Figure S2b.

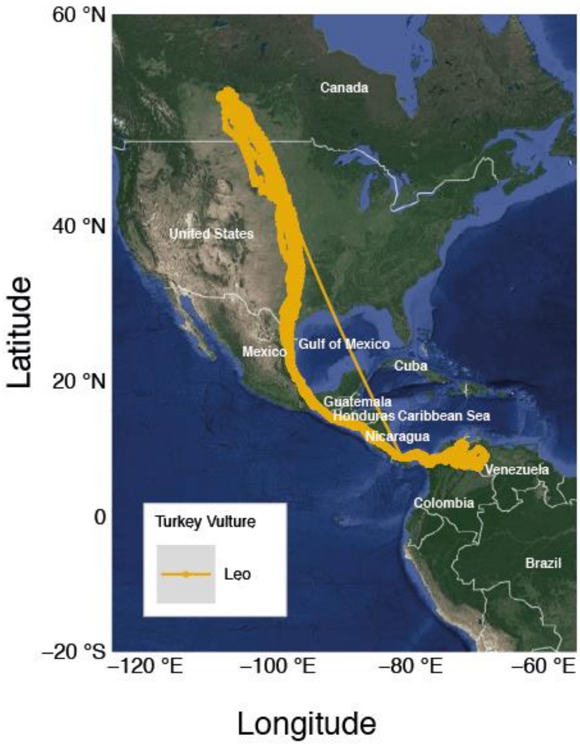

Figure S2b: Movement trajectory of Leo

From the raw trajectory a movement parameter profile (i.e. speed in this case) is computed, serving as the input to the DWT analysis (Figure S2c). The DWT is flexible regarding the input movement signal and different movement metrics can be used for the analysis (e.g. acceleration, turning angles). Based on biological domain knowledge, movement speed seems a reasonable choice to distinguish between migratory versus non-migratory phases and therefore was selected here as the movement signal.

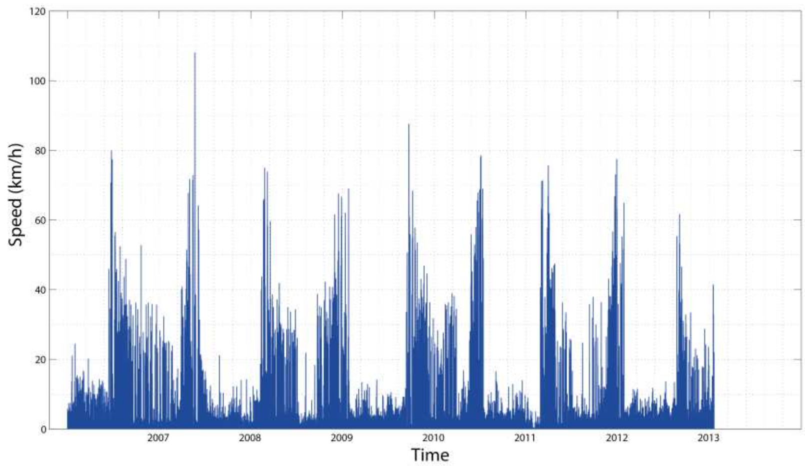

Figure S2c: Speed profile of Leo

In the next step, the movement signal is decomposed through the discrete wavelet transform. To do so, a mother wavelet function needs to be chosen. The family of Daubechies are a well-known class of wavelet function and the order 4 (db4) was chosen according to similar application of DWT in the literature. Other mother wavelet functions are available (e.g. Haar, Morlet, Mexican hat) and their application depends on the shape of the response one is interested in (Boggess & Narcowich 2009). For instance, certain wavelet functions are able to resolve small, abrupt discontinuities, whereas others may better capture linear changes in the movement. In general, we recommended to use db4 for the application of our method. The outputs of the DWT decomposition are two sets of approximation and detail sub-bands, respectively, shown in Figure S2d.

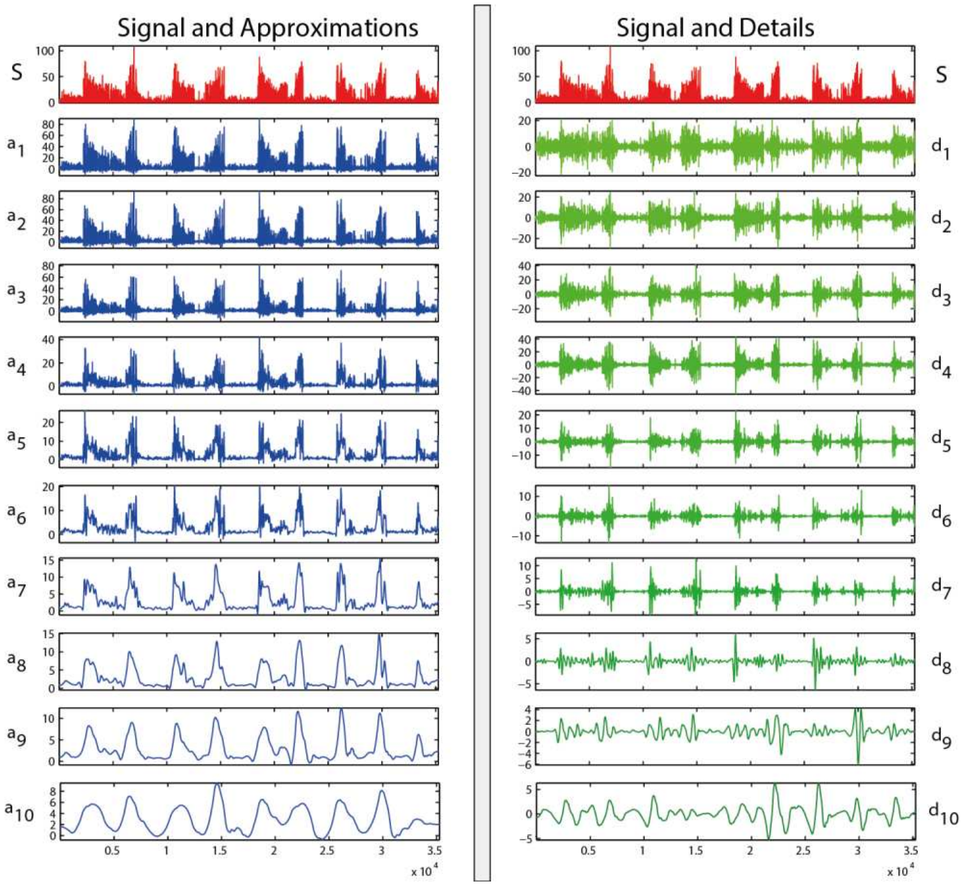

Figure S2d: Illustration of the approximation and detail sub-bands resulting from the DWT transform

While the DWT segmentation method is taking advantage of both of these components, each is treated differently and has different analysis purposes, i.e. approximation sub-bands for detection of behavioural segments and detail sub-bands for detection of change points. As the use of these two components is independent, different parameterizations are needed.

2.1. Behavioural segmentation using the approximation sub-bands

Visual inspection of the different approximation sub-bands shows that Level 8 can be an appropriate level in this case because peaks are adequately distinguished from the rest of the signal indicating periodic patterns. A selection of the sub-bands are shown in Figure S2e: at the initial levels of decomposition, the localization effects are not strong enough yet to suggest differences between migration phases, which also make it impossible to apply any peak analysis. However, the localization effect is becoming stronger with increasing levels, which can be clearly seen in the resulting peaks versus flat regions at the Levels 7, 8 and 9. Therefore, the goal is to reach a level where segments related to the same behaviour are expressed in peaks with similar height and widths. Accordingly, level 8 was finally chosen because distinct phases became apparent as the level where peak analysis can potentially distinguish between different segments.

After the approximation level is set, the next parameter that needs to be set is the peak height to distinguish between different phases (blue triangles in Figure S2e). As can be seen at the 8th level, the height threshold is chosen in such a way that the widths of the resulting peaks (determined through the peak prominence shown as orange lines in Figure S2e) are in agreement with the length (i.e. duration) of migratory patterns. As can be seen in Figure S2e, the peaks related to migratory seasons are well distinguished from the rest of the signal and the combination of the height and the width of the peaks will help to identify the segments. Applying the same threshold on the 7th or 9th level (Supplementary Material S6) results in similar segments indicating that the method is fairly robust to the choice of decomposition level.

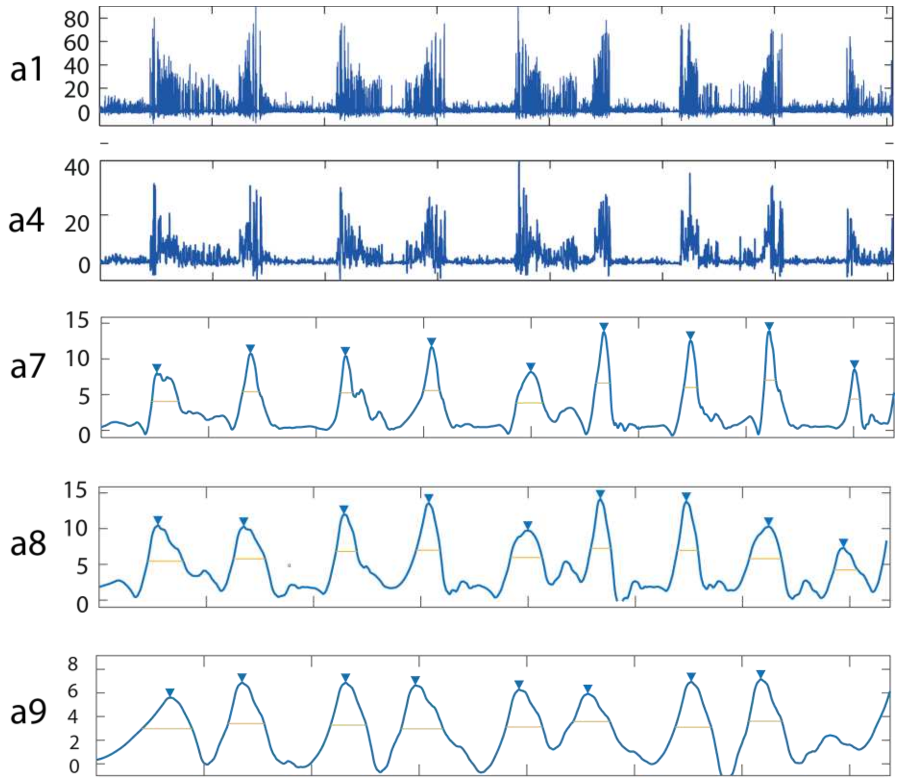

Figure S2e: Parameterization of peak analysis in the approximation sub-band. The height of the peak are shown as blue triangles and the peak widths as horizontal orange lines.

The user is recommended to carefully inspect the variation of the resulting peaks and choose the level, as well as peak height, where he/she thinks that it would work best to distinguish among different phases in the data. A sensitivity analysis of other levels, same as what is done in Supplementary Material S6 is also recommended to assure the robustness of the method to the choice of peak height(s).

2.2. Change point analysis using the detail sub-bands

Since the detail sub-bands contain the high-frequency information, the analysis based on these sub-bands focuses on the detection of fine-scale changes in movement. Same as with the approximation, first the relevant level for the detail sub-bands needs to be selected. This is also done by visual inspection of the detail sub-bands and further informed by previous knowledge regarding at which level biologically-driven changes may occur. As shown in Figure S2f, the frequency resolution at the initial levels of decomposition (e.g. d1) is so high that a distinction between different phases is difficult. Then, at Level 5 one can start to observe the different regimes of values for migratory and non-migratory phases, as well as some finer-scale changes within these phases. Therefore level 5 is selected for further analysis.

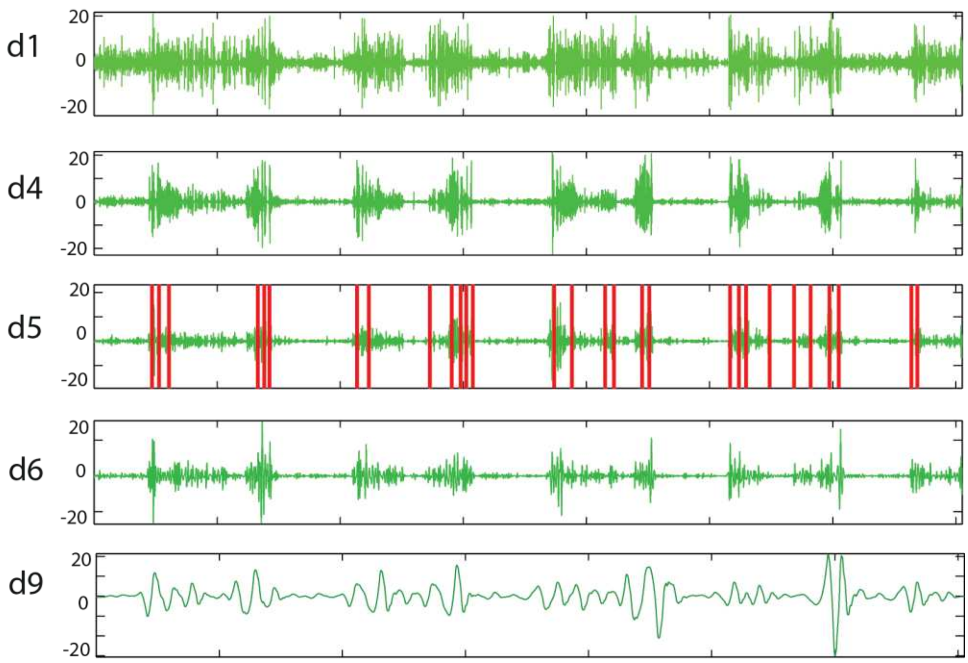

Figure S2f: Parameterization of change point analysis in the detail sub-bands. For identifying change points, the difference between detail coefficients (D_t – D_{t+1}) are computed for every point and the points exceeding a certain threshold are identified as change points.

The next step is to identify change points by setting a threshold for the difference between the wavelet coefficients (vertical red bars in Figure S2f). The value of this threshold also depends on the data used. For example in this case, the detail coefficients in the non-migratory season are around zero at the selected Level 5, whereas the values in the migratory seasons are much higher. Therefore a value of 1.0 was selected to identify the change points, that is, every point whose coefficient value is 1.0 units higher than the value of its preceding point.

The results after the thresholding are shown in Figure S2g (middle panel). It can be seen that numerous change points have been identified (in this case a total 162).

An optional concatenation step to obtain fine scale behavioural segments can be applied. This makes it possible to better match the results to the expected duration of the patterns that are sought (i.e. migratory patterns). In our case, all the segments with a length shorter than 500 fixes (i.e. corresponding to ~20 days) were concatenated to their preceding segment. The results are shown in Figure S2g (bottom). Here, one can observe that all the migratory seasons are recovered, with some finer-scale segments in some of the non-migratory periods remaining. These results can be used for further investigation of the species behaviour at finer scales.

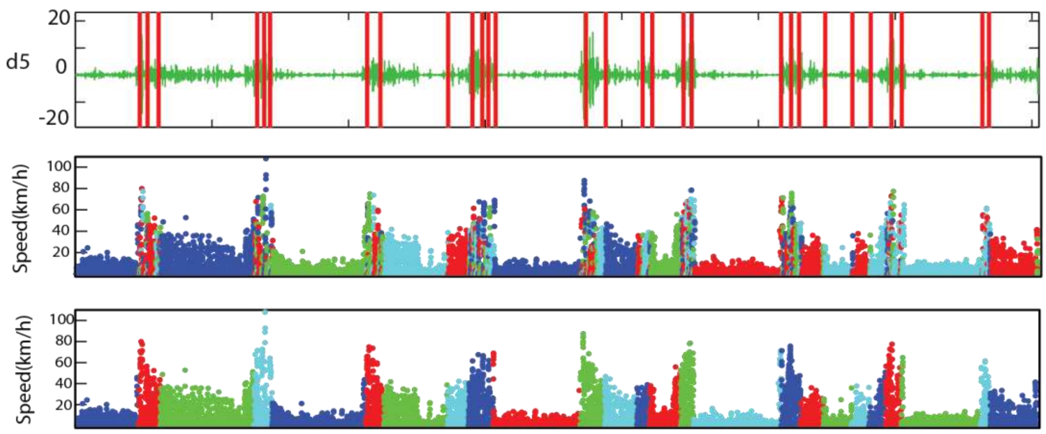

Figure S2g: Resulting segments identified through the change points (middle) and after concatenation of short segments (bottom).

Same as for the approximation sub-bands, the user is recommended to carefully inspect the variation of the resulting detail sub-bands and choose the level where he/she hypothesizes that different phases in the data are consistent with biological knowledge of the system. Setting the threshold for identifying change points is also dependent on the extent to which finer-scale changes are to be recovered. The concatenation step is totally optional, with the length of segments for concatenation be specified by considering a minimum duration for the patterns of interest. That is in our case, a migration period cannot be shorter than 20 days. Therefore 500 fixes (indicating 20 days) were considered as the minimum length where all the migratory patterns as well as some additional segments were identified.

Appendix S3

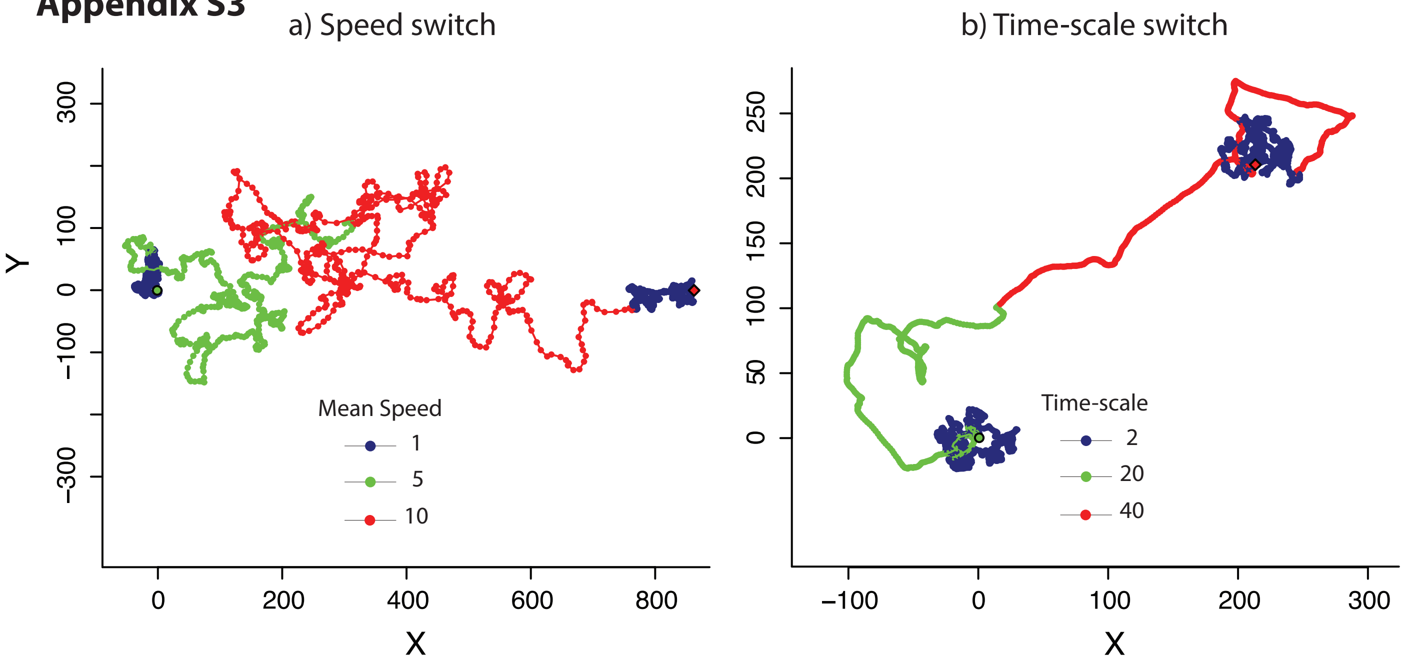

Figure S3a. Examples of simulated tracks used for validation: a) Speed switch model, where the values of mean speed are changing in different segments (1,5,10,1). b) Time-scale switch, where changing the time-scale (2,20,40,2) for the four modes results in segments of different tortuosity.

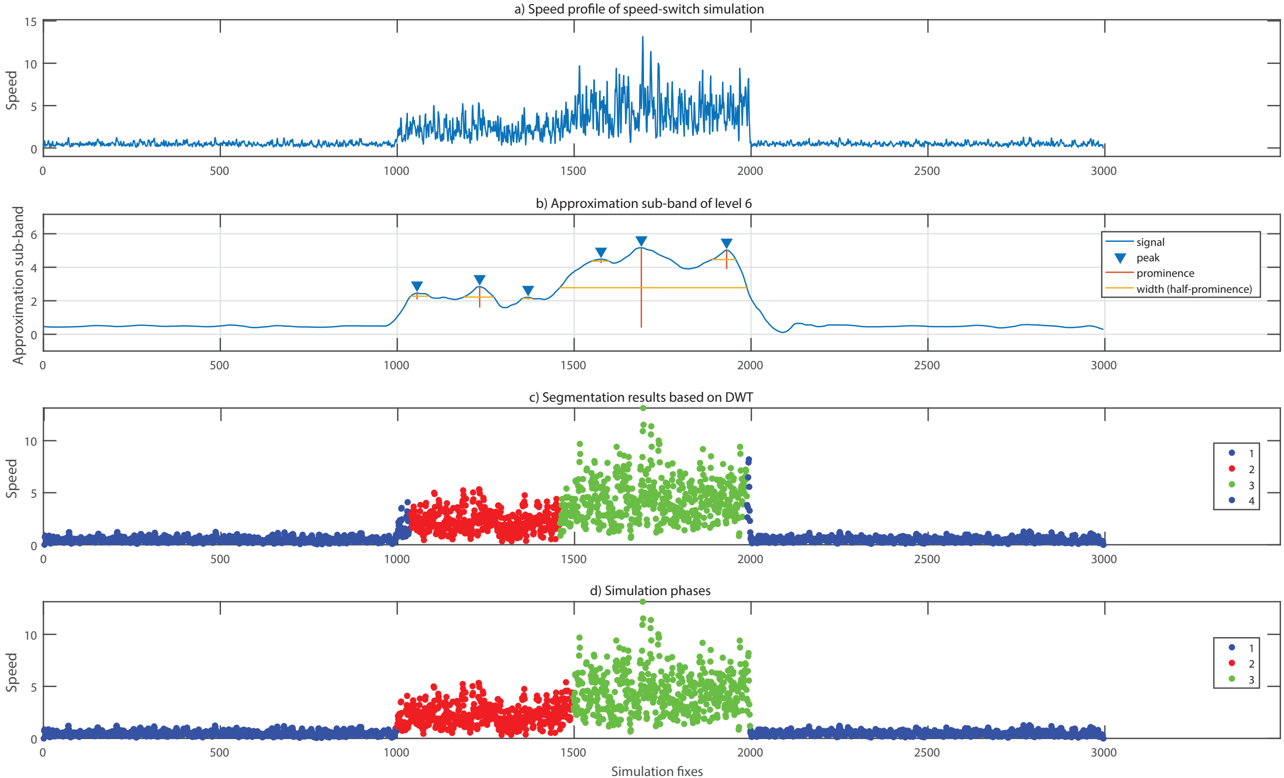

Figure S3b. Applying the proposed segmentation method on the speed-switch simulation. a) Speed profile as the input signal for wavelet analysis. b) Detected peaks in approximation level 6 by thresholding the height of the peaks, in order to distinguish between the four phases. c) Segmentation results based on the width of the extracted peaks. The resulting 4 segments are closely representing the 4 phases in the simulation (shown in d).

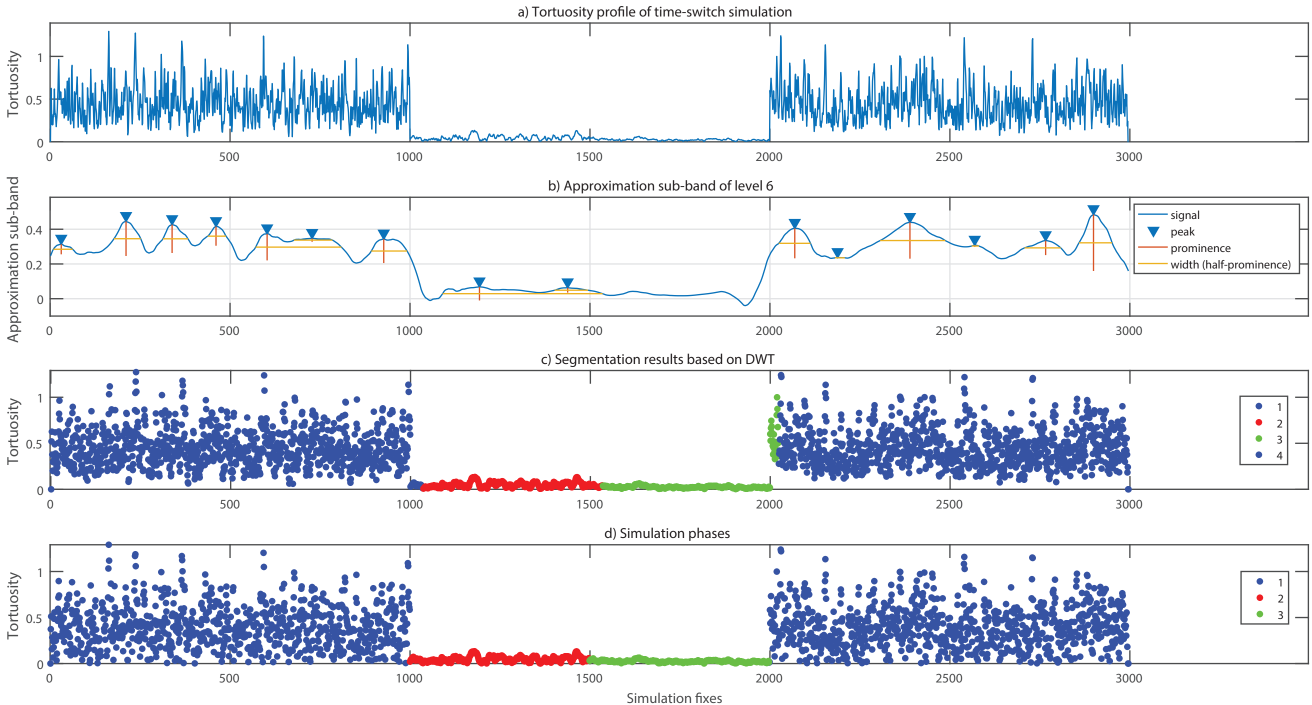

Figure S3c. Applying the proposed segmentation method on the time-switch simulation. a) Tortuosity profile as the input signal for wavelet analysis. b) Detected peaks in approximation level 6 by thresholding the height of the peaks, in order to distinguish between the four phases. c) Segmentation results based on the width of the extracted peaks. The resulting 4 segments are closely representing the 4 phases in the simulation (shown in d).

Comparison of the DWT-based method to the state-of-the-art segmentation methods to detect continuous transitions in movement

In order to demonstrate the applicability of the method to detect continuous transitions in movement (i.e. cases where behaviour is in flux), we generated a track based on the speed-switch simulation presented in Gurarie et al. (2016). We considered 5 phases (each of 1000 fixes) for the simulation, where the speed values were increasing in a step-wise, incremental fashion from 1 to 5, i.e. the average speed in the first phase was 1, increasing in increments of 1 until the top speed of 5 is reached in the fifth phase. The simulated track is shown in Figure S4a.

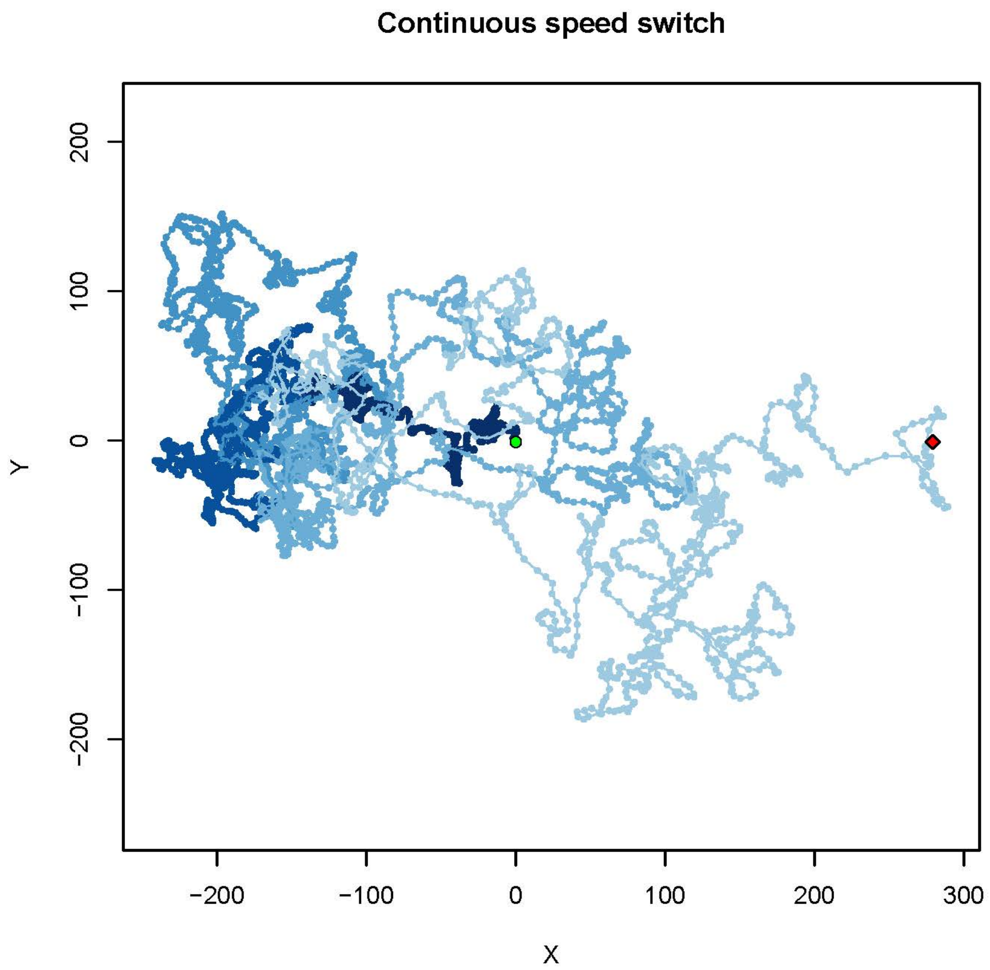

Figure S4a: The simulated track with continuous increase in the speed values from 1 to 5 in five phases.

Whereas these increments do not represent perfectly smooth transitions, they nevertheless are clearly smoother than the marked speed switches in the original simulation by Gurarie et al. (2016). We then tested the capability of state-of-the-art methods, i.e. FPT, BPMM and BCPA to identify such changes and compared to the performance of the DWT method. The parameters of the methods were set to the same values as the speed-switch simulation in Gurarie et al. (2016).

In the first case, FPT did not show any flexibility and resulted in completely missing the segments. The result is shown in Figure S4b.

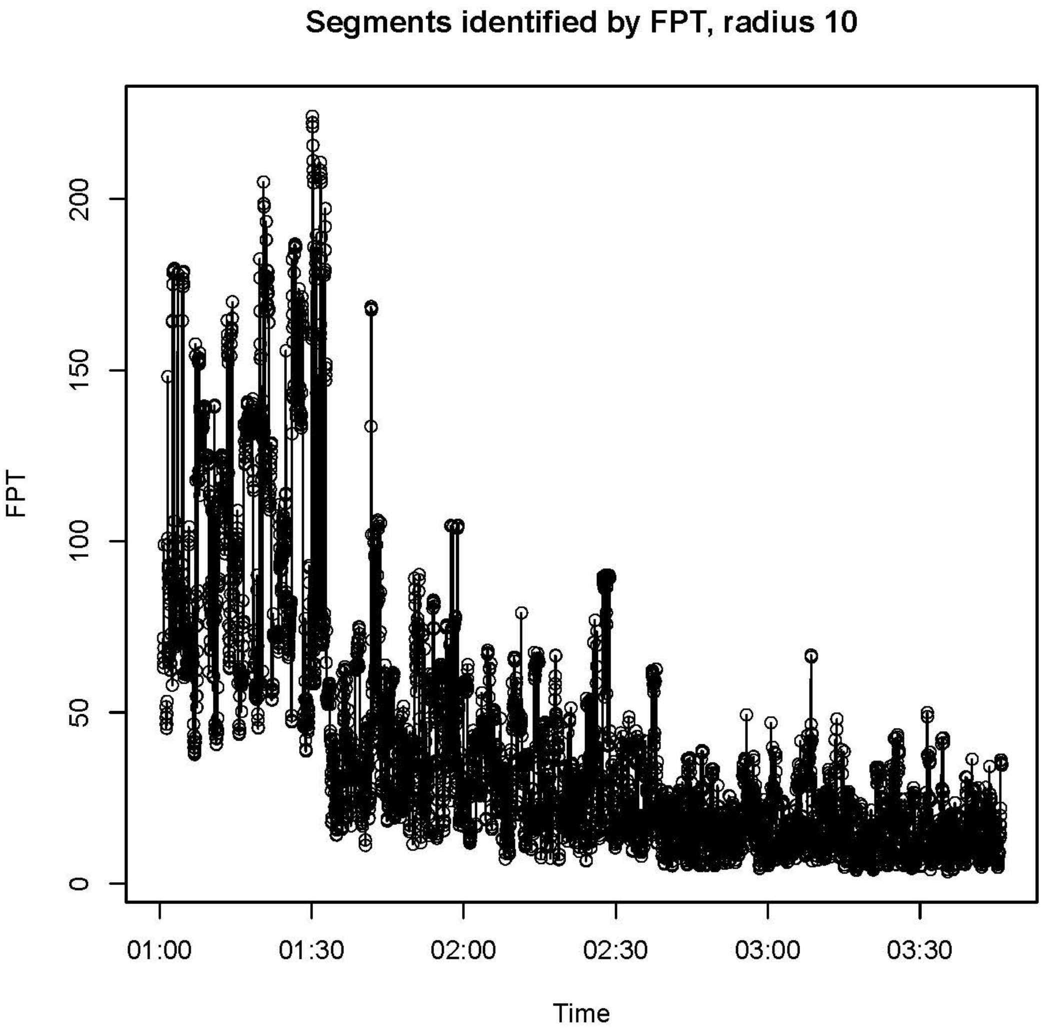

Figure S4b: Segmentation results based on FPT. The method was not able to identify the continuous changes

The result of applying BPMM is shown in Figure S4c. Although some segments are detected, the method fails, detecting far too many segments (i.e. 30). This may be mainly due to the sensitivity of the method to abrupt variations, i.e. excessive change points are detected not only in the changes from one phase to the other, but also within the phases.

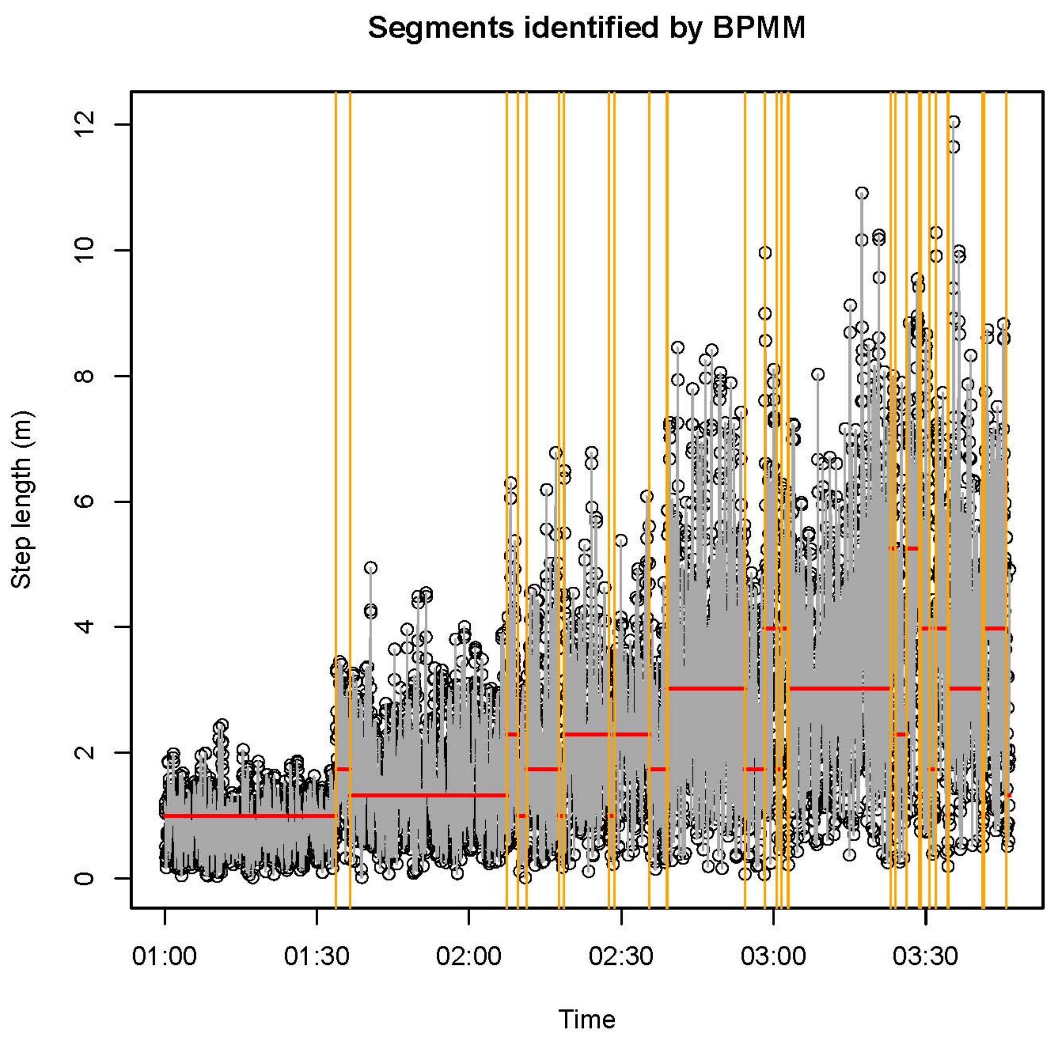

Figure S4c: Segmentation based on BPMM resulted in 30 segments. Since the autocorrelation effects are not considered, the method over-splitted the trajectory especially in cases where the variations in the data were higher (phases 3 to 5)

The result of applying BCPA is shown in Figure S4d. While BCPA detected less segments (i.e. 11), the positions of the extracted segments are not representative.

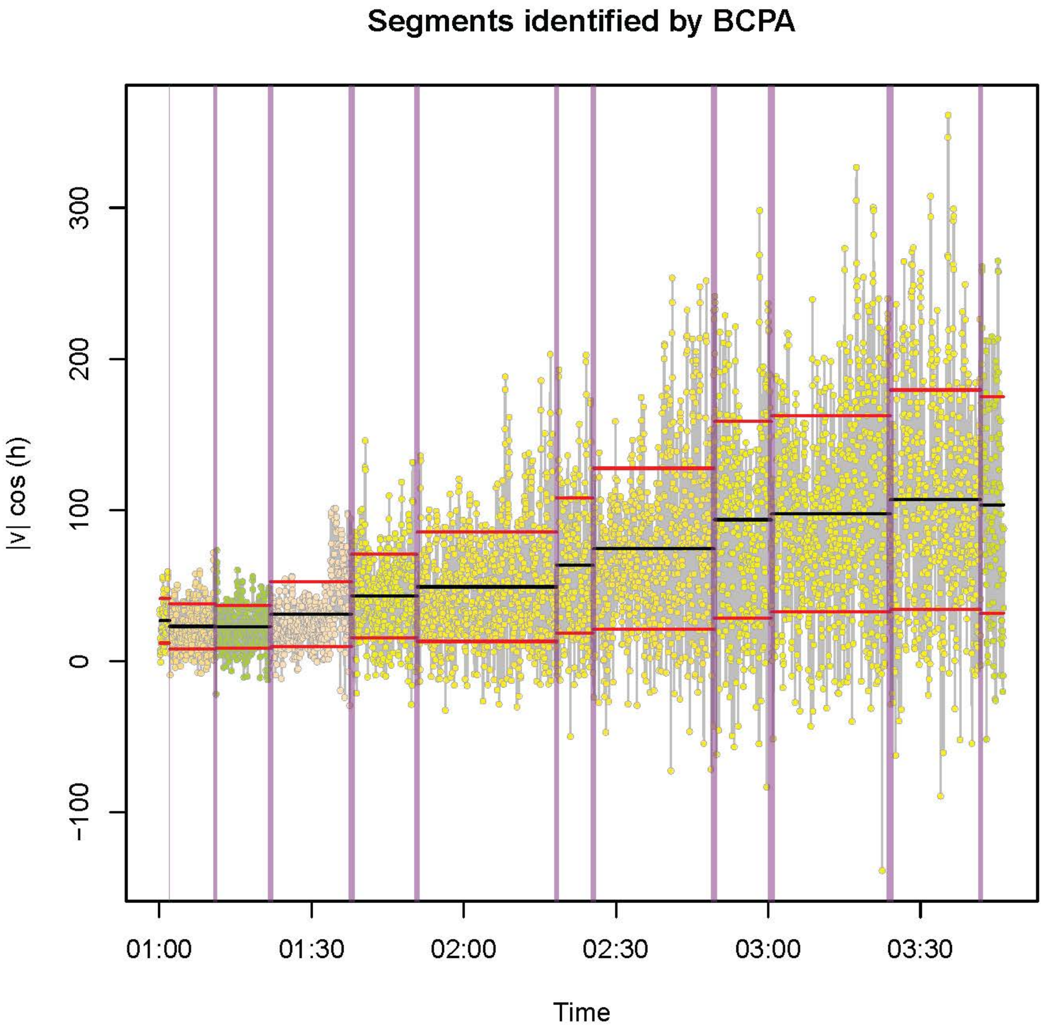

Figure S4d: Segmentation results based on BCPA. The performance slightly improved by extracting 11 segments, however redundant segments and incorrect positioning of change points are obvious.

Finally, the proposed DWT method was applied to the continuous simulation and it successfully retrieved the 5 segments. Although the selection of peak heights needs to be done in a careful manner, the method performed better than the other three methods. This shows that the proposed method not only can detect discrete forms of changes, but also enables identifying continuous changes in movement.

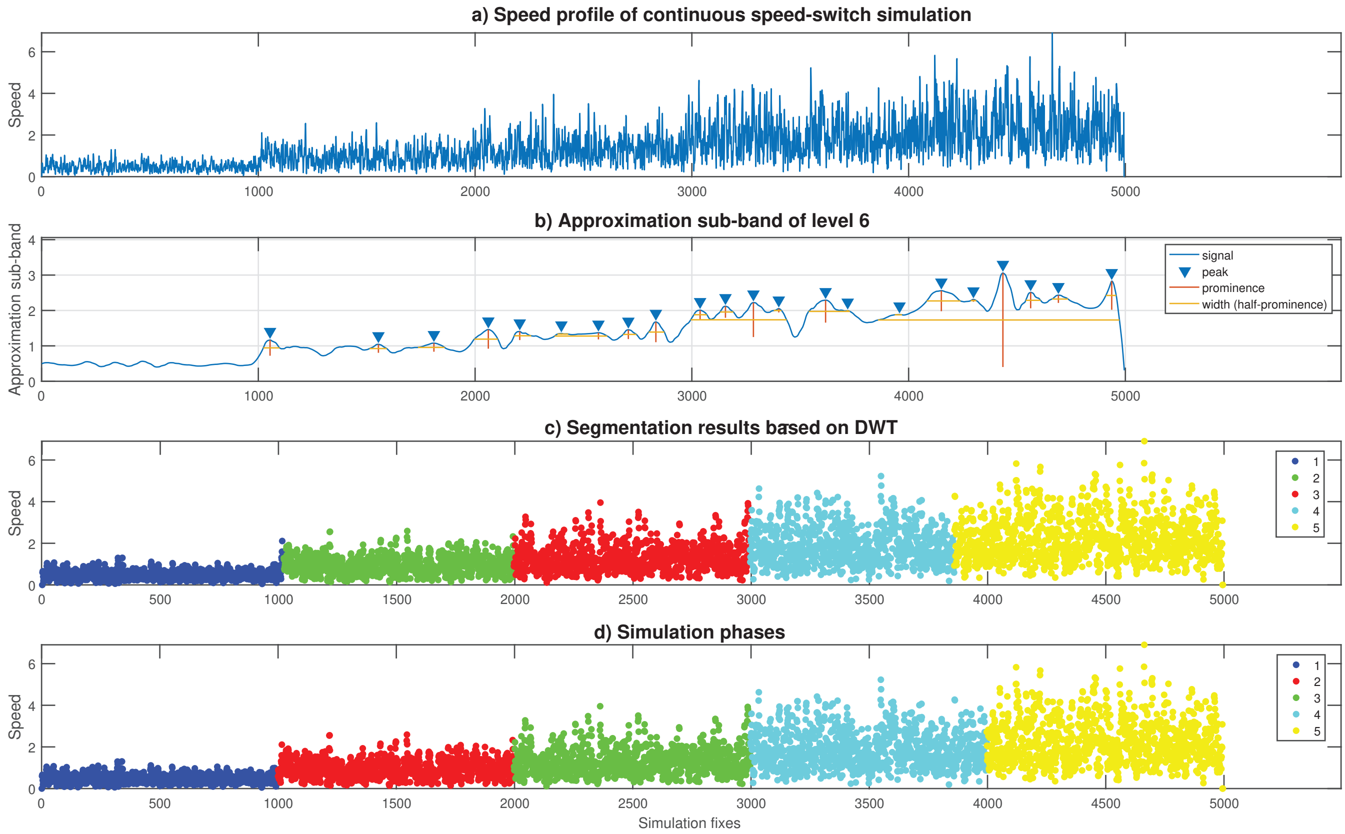

Figure S3e: Segmentation results based on DWT. The method detected exactly 5 segments. Also, the segments were matching best to the annotations compared to the other methods.

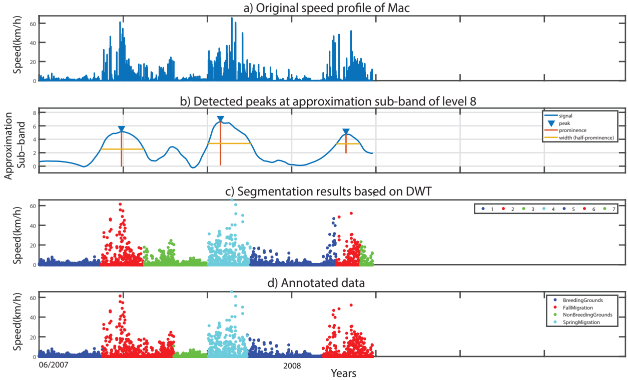

Figure S5a. Applying the proposed segmentation method on data of Mac. a) Speed profile of Mac as the input signal for wavelet analysis. b) Detected peaks in approximation level 8 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 7 segments are closely representing the annotated data (shown in d).

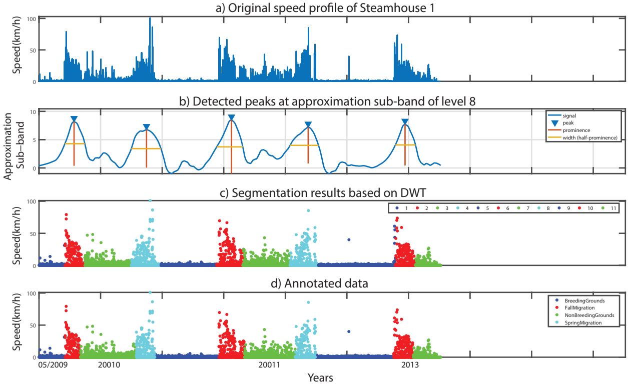

Figure S5b. Applying the proposed segmentation method on data of Steamhouse 1. a) Speed profile of Steamhouse 1 as the input signal for wavelet analysis. b) Detected peaks in approximation level 8 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 11segments are closely representing the annotated data (shown in d).

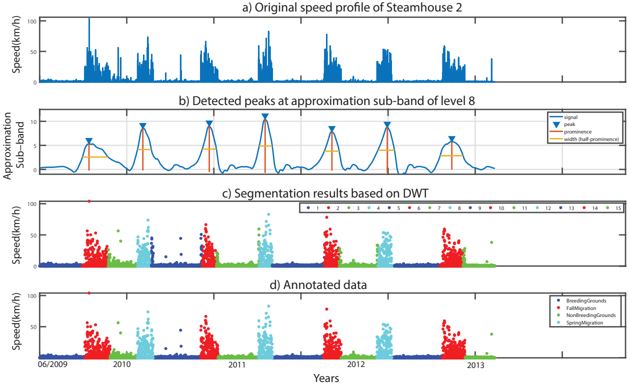

Figure S5c. Applying the proposed segmentation method on data of Steamhouse 2. a) Speed profile of Steamhouse 2 as the input signal for wavelet analysis. b) Detected peaks in approximation level 8 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 15 segments are closely representing the annotated data (shown in d).

Appendix S6

Sensitivity analysis of DWT segmentation method

In the following, we show that the results of employing the DWT method on the 7^{th} and 9^{th} decomposition levels remain comparable to the results of the 8^{th} decomposition level. This was done in order to demonstrate the robustness of the method to the choice of decomposition level and that the choices stay robust as long as an approximately optimal level is chosen.

First, we show the results for the 7^{th} level for the four turkey vulture individuals, followed by box plots to show the difference to the annotations (FigureS6a-e). The results show that DWT performed the same as for the 8^{th} level in terms of recovering the number of segments. The average difference to the annotation remained the same as for the 8^{th} level, i.e. 7 days.

Second, the method was applied on the 9^{th} level (FigureS6f-j). Since the signal has been smoothed further, the method resulted in missing some the segments in some of the individuals. The average difference also increased to 14, however result didn't completely deteriorate compared to the original 7 days of the 8^{th} level.

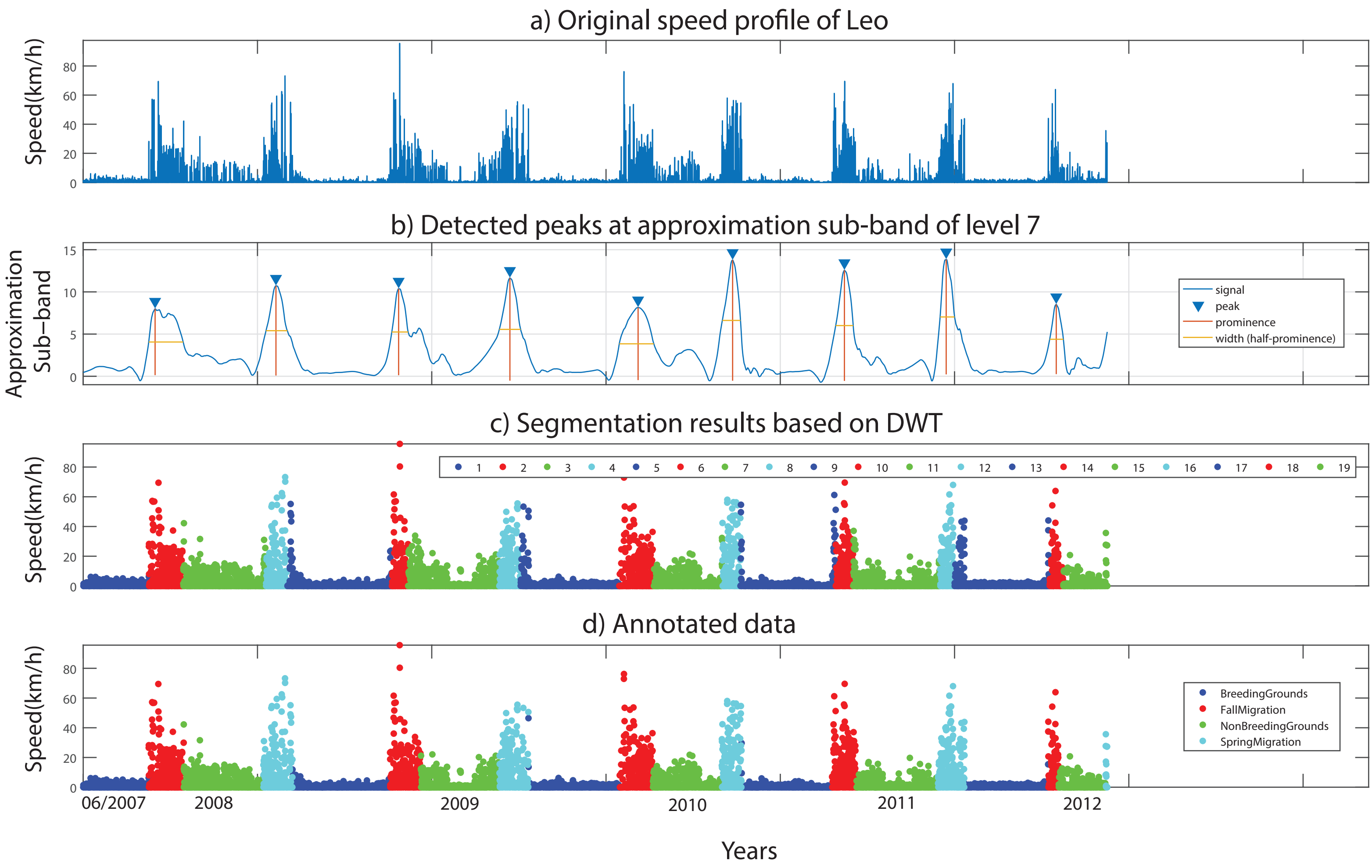

Figure S6a. Applying the proposed segmentation method on data of Leo. a) Speed profile of Leo as the input signal for wavelet analysis. b) Detected peaks at approximation level 7 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 19 segments are closely representing the annotated data (shown in d).

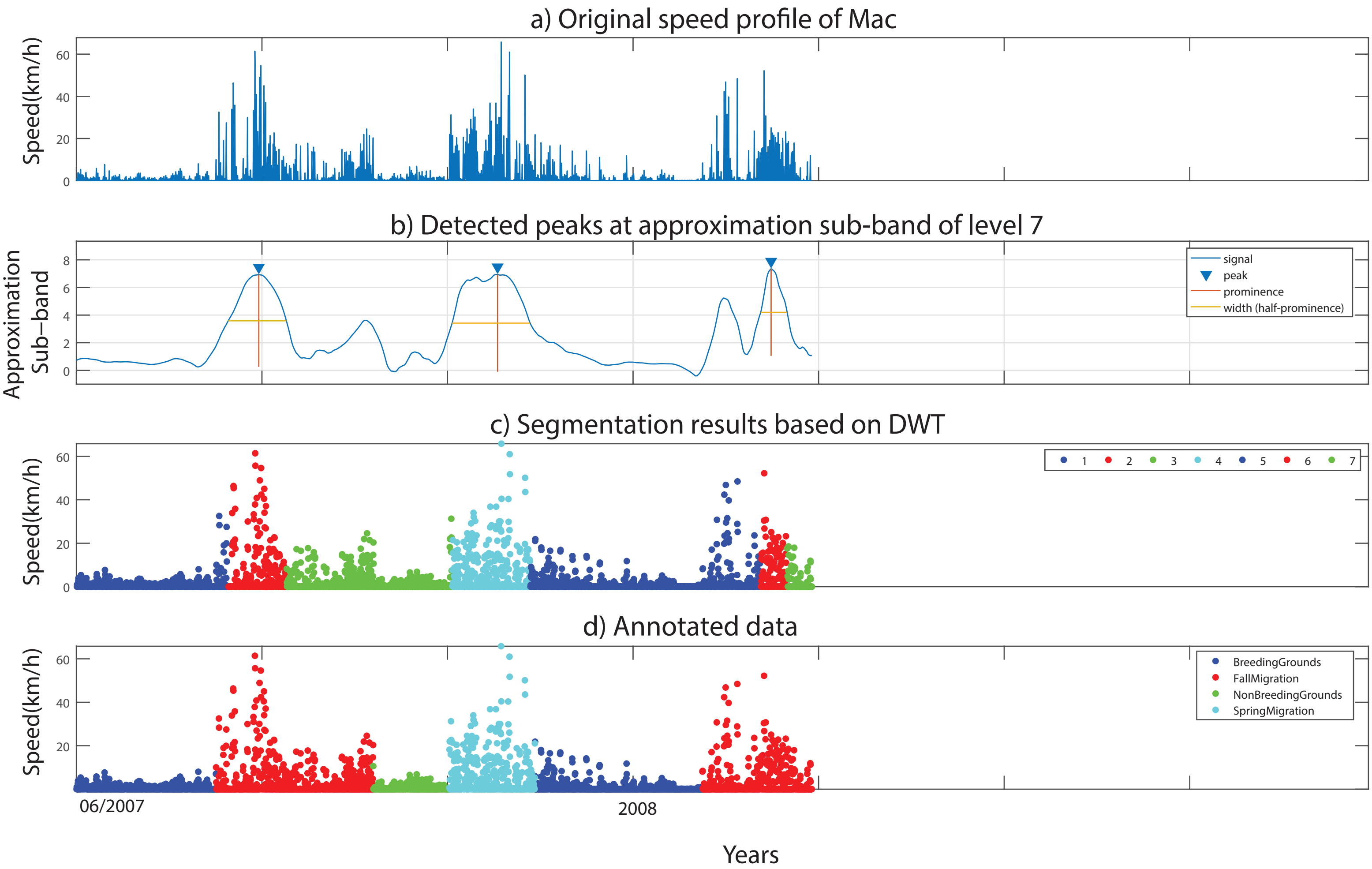

Figure S6b. Applying the proposed segmentation method on data of Mac. a) Speed profile of Mac as the input signal for wavelet analysis. b) Detected peaks at approximation level 7 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 7 segments are closely representing the annotated data (shown in d).

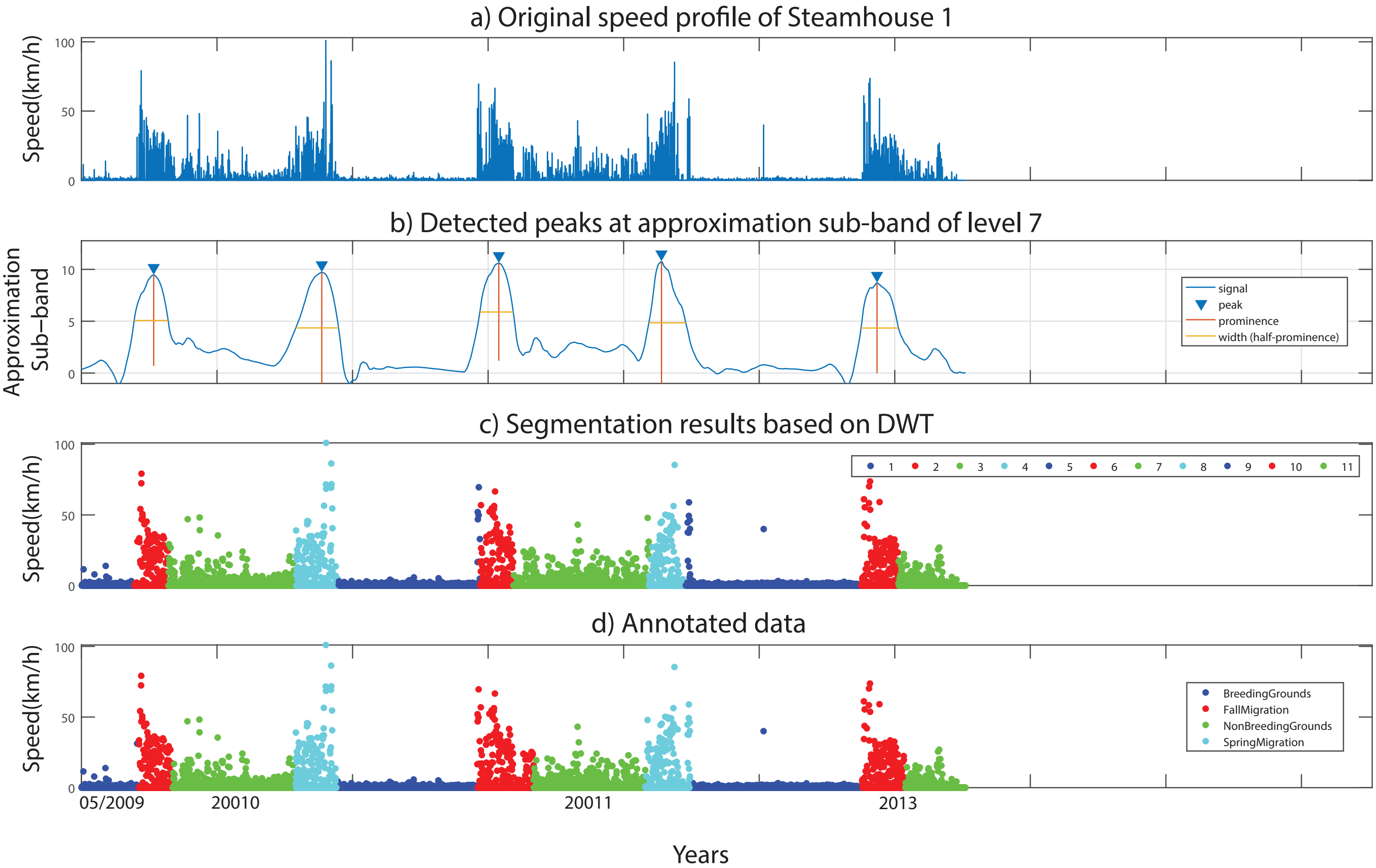

Figure S6c. Applying the proposed segmentation method on data of Steamhouse 1. a) Speed profile of Steamhouse 1 as the input signal for wavelet analysis. b) Detected peaks at approximation level 7 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 11segments are closely representing the annotated data (shown in d).

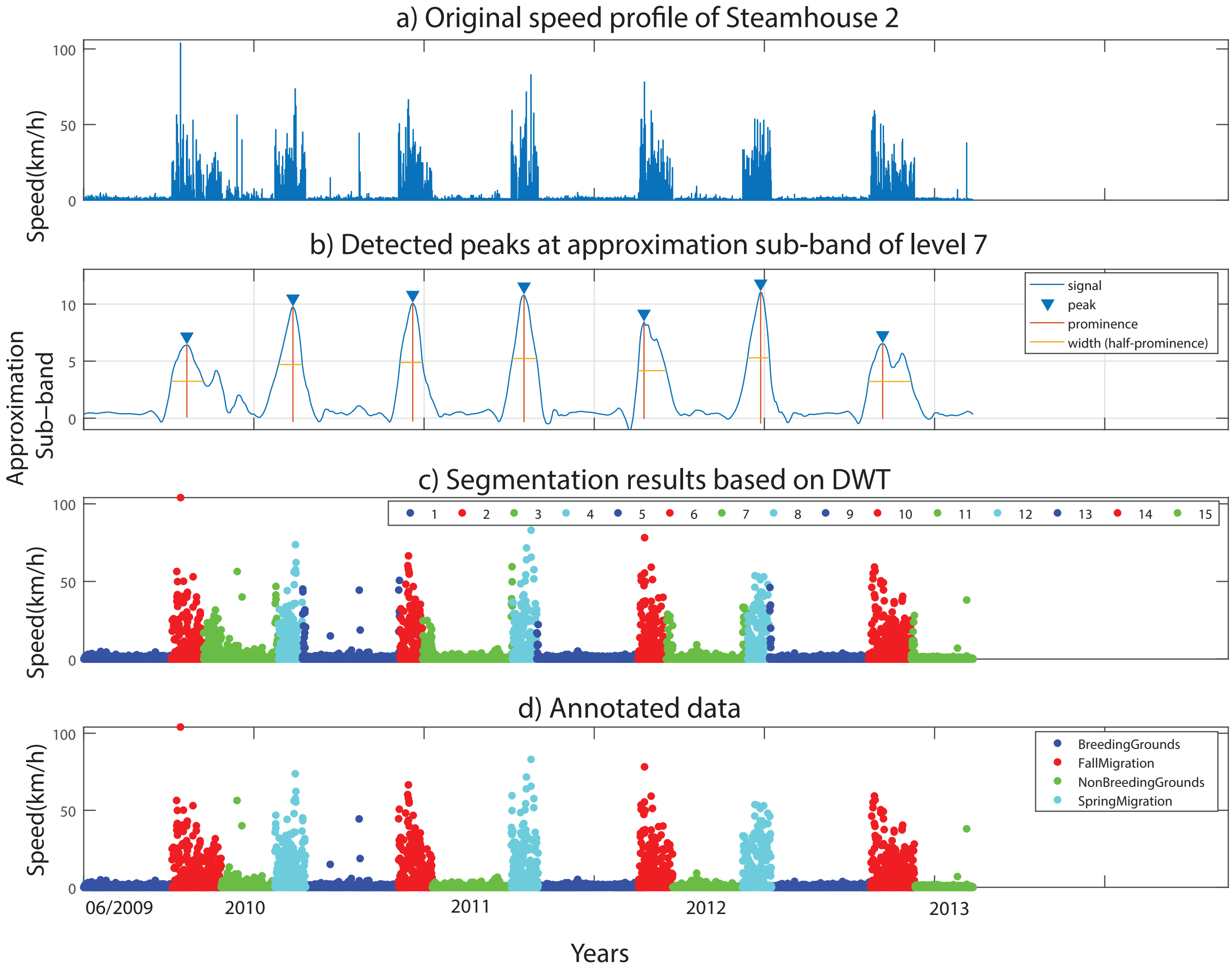

Figure S6d. Applying the proposed segmentation method on data of Steamhouse 2. a) Speed profile of Steamhouse 2 as the input signal for wavelet analysis. b) Detected peaks at approximation level 7 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 15 segments are closely representing the annotated data (shown in d).

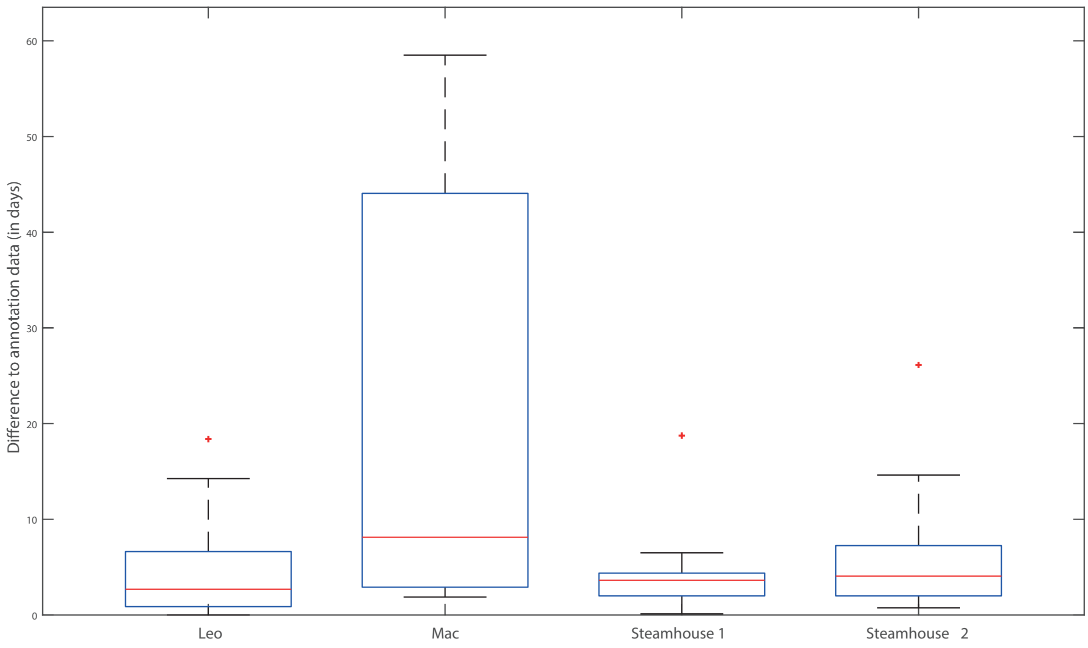

Figure S6e. Temporal difference between the extracted segments and annotated segments at the 7th level.

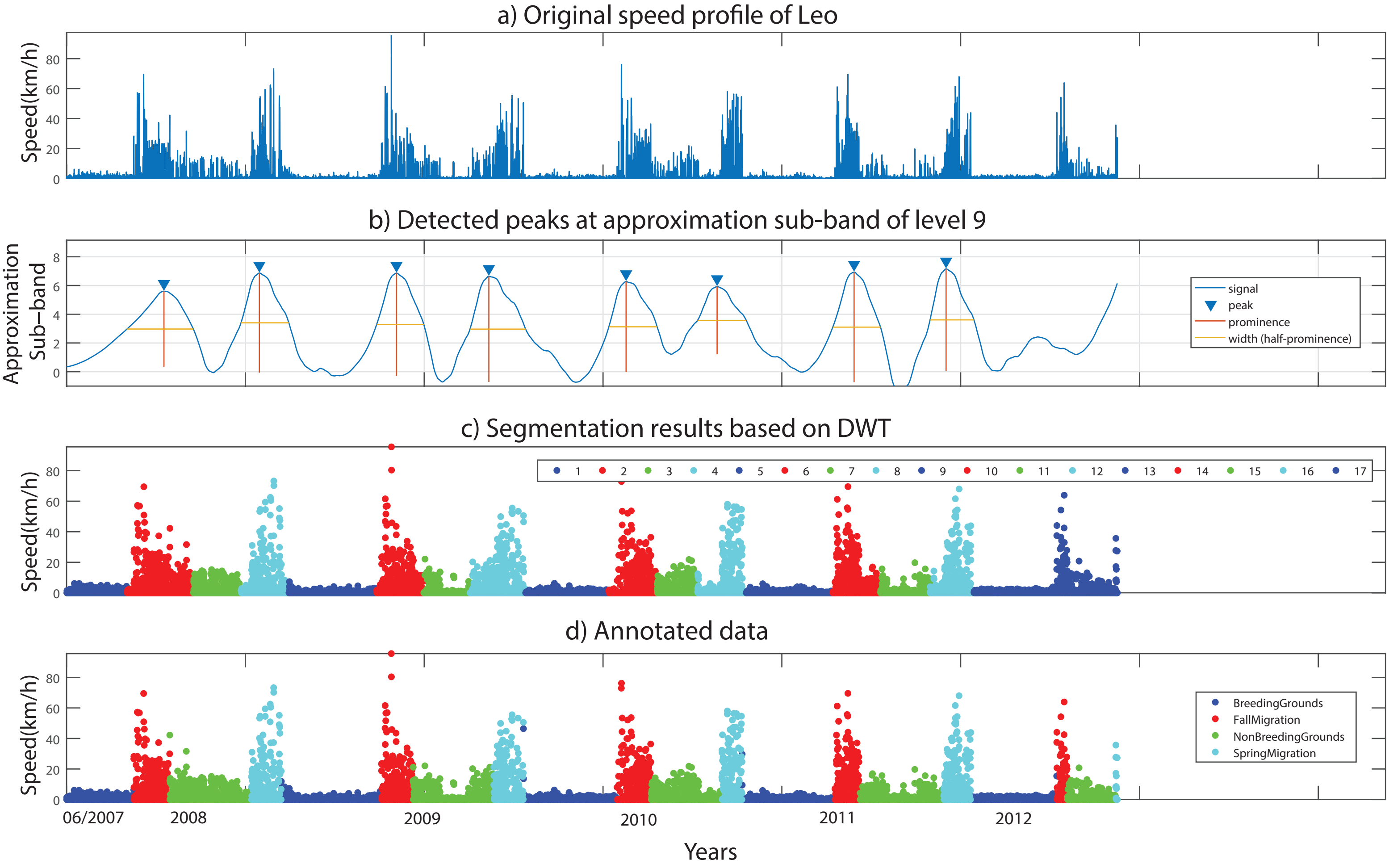

Figure S6f. Applying the proposed segmentation method on data of Leo. a) Speed profile of Leo as the input signal for wavelet analysis. b) Detected peaks at approximation level 9 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. Most of the resulting 17 segments represent the annotations, however some segments are missed due to the signal getting further smoothed at the 9th level.

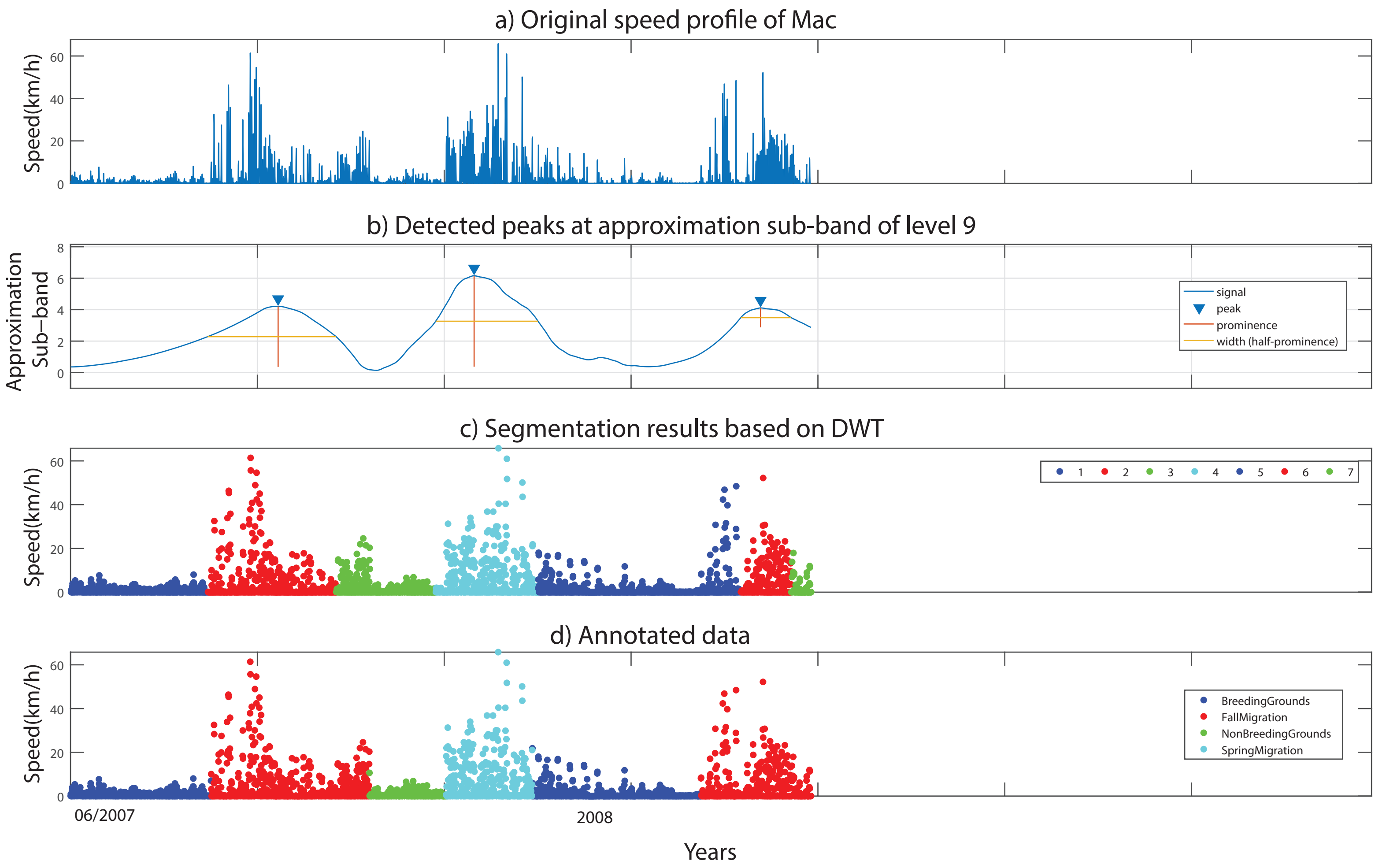

Figure S6g. Applying the proposed segmentation method on data of Mac. a) Speed profile of Mac as the input signal for wavelet analysis. b) Detected peaks at approximation level 9 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. Most of the resulting 7 segments represent the annotations, however some segments are missed due to the signal getting further smoothed at the 9th level.

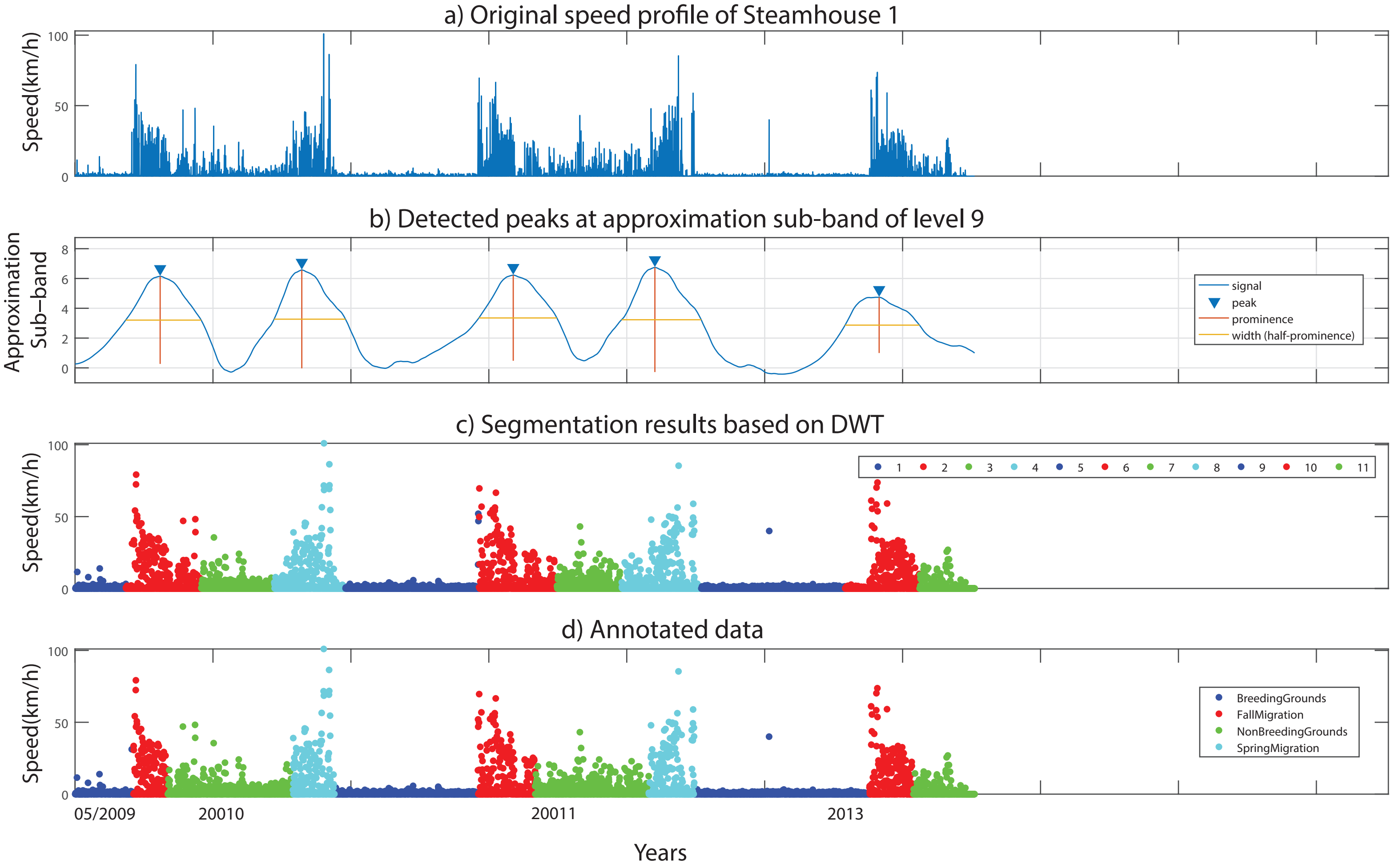

Figure S6h. Applying the proposed segmentation method on data of Steamhouse 1. a) Speed profile of Steamhouse 1 as the input signal for wavelet analysis. b) Detected peaks at approximation level 9 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 11 segments represent the annotations, however the peak widths are getting wider due to the signal getting further smoothed at the 9th level, resulting in wider segments.

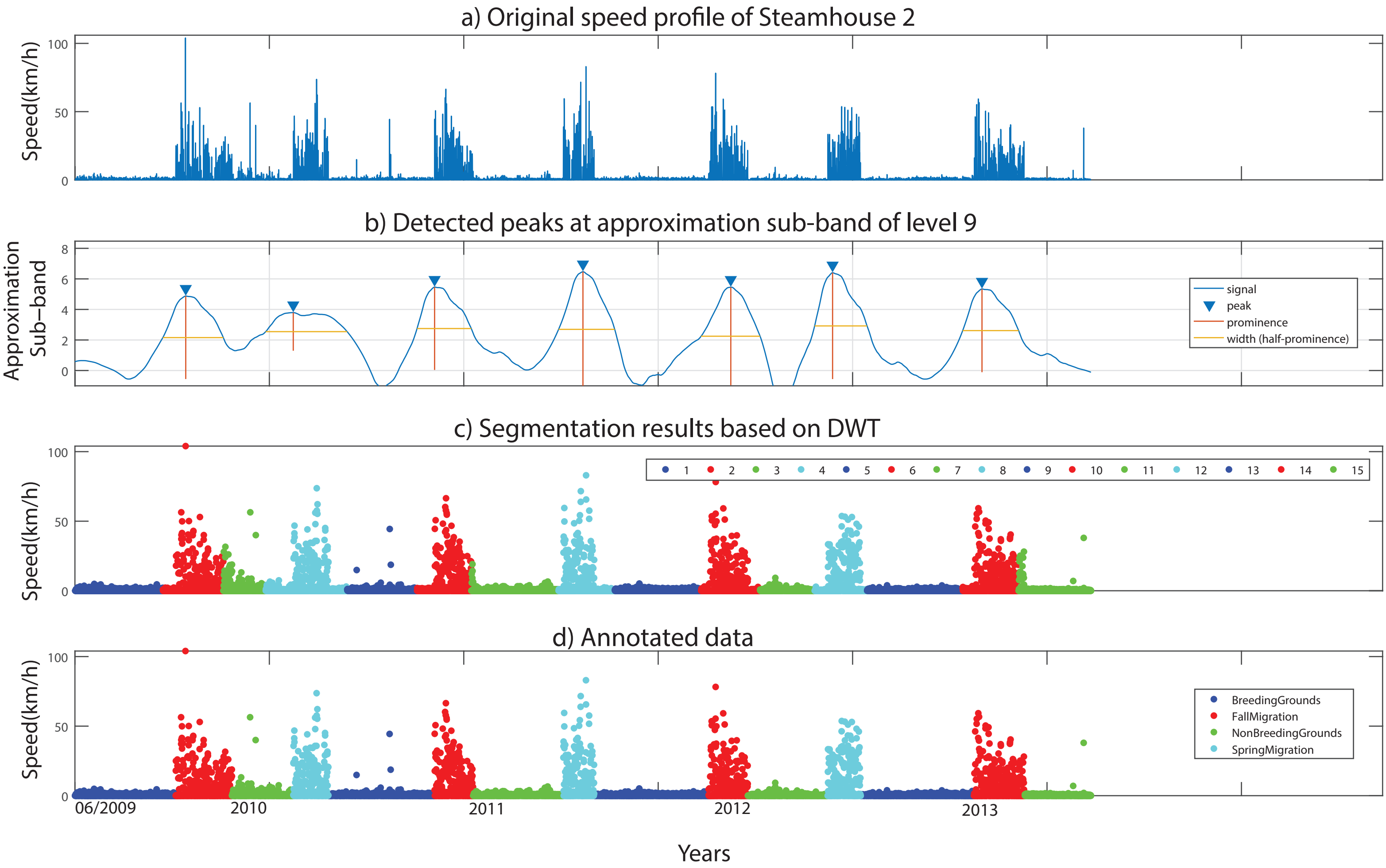

Figure S6i. Applying the proposed segmentation method on data of Steamhouse 2. a) Speed profile of Steamhouse 2 as the input signal for wavelet analysis. b) Detected peaks at approximation level 9 by thresholding the height of the peaks, in order to distinguish between migratory and non-migratory seasons. c) Segmentation results based on the width of the extracted peaks. The resulting 15 segments represent the annotations, however the peak widths are getting wider due to the signal getting further smoothed at the 9th level, resulting in wider segments.

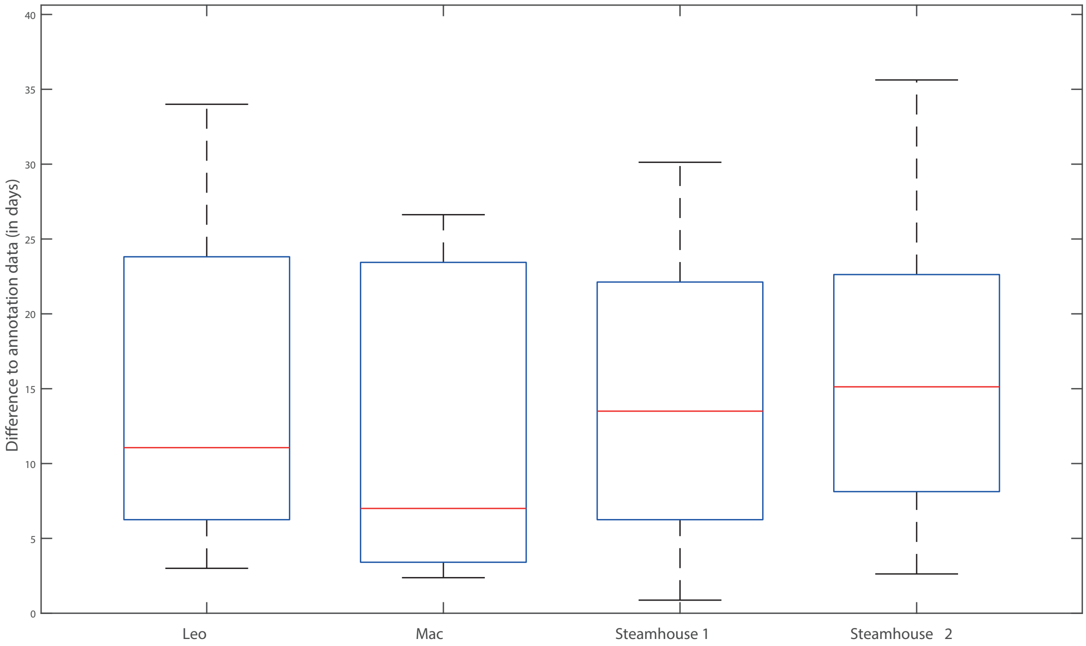

Figure S6j. Temporal difference between the extracted segments and annotated segments at the 9th level.